

Performance and Quality of Dementia Prediction Models Using Electronic Health Records: A Systematic Review

National Centre
for Healthy Ageing

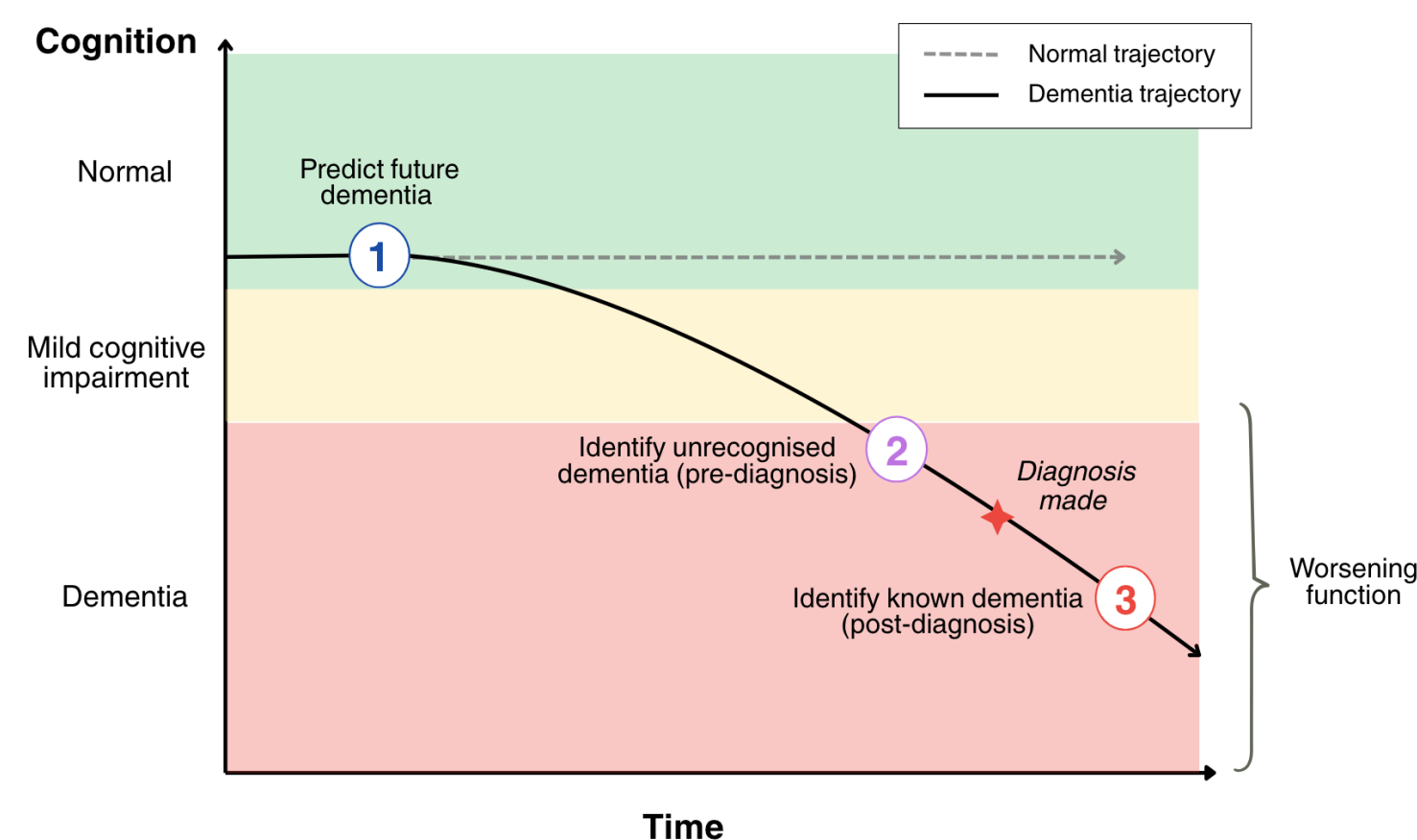
A partnership between
MONASH University | Peninsula Health

Alicia Lu,^{1,2} Velandai Srikanth,^{1,2,3} Sarah Westworth,^{1,2} Yue-Guang Baey,^{4,5} Chris Moran,^{1,5,6} Richard Beare,^{1,2,7} Kristy Siostrom,^{1,3} Nadine Andrew,^{1,2} Taya Collyer,^{1,2}

Background

- Many people with dementia remain unrecognised, delaying diagnosis and care
- Prediction models using routine electronic health record (EHR) data offer a scalable opportunity to identifying dementia across the disease trajectory
- However, the quality, performance, and readiness of existing EHR-based models remain unclear**

Prediction model tasks along the dementia trajectory



Methods

- ✓ Systematic review of EHR-based dementia prediction models
- ✓ **Databases:** Medline, EMBASE, Scopus, IEEE Xplore, ACM, grey literature sources (Google scholar, preprint servers)
- ✓ **Search period:** inception to July 2024
- ✓ Data extraction using CHARMS checklist
- ✓ Risk of bias assessment using PROBAST

Current Research Landscape

11,206 records screened → 56 studies included
(describing 434 models and 155 external validations)

80% of models
used structured
EHR data

34% of models
were diagnostic

71% of models
were developed in
the United States

11% of models
were externally
validated

Key Findings

1 Model prediction tasks were often poorly defined

The intended prediction task was unclear in 63% of studies.

Many failed to distinguish between predicting future dementia, identifying unrecognised dementia, and identifying existing dementia.

2 Reference standards were often unreliable

Only 4% of models used established clinical diagnostic criteria as the reference standard.

Most relied on routinely coded diagnoses, despite known misclassification of dementia in healthcare records.

3 Model evaluation was often incomplete

Most models focused on discrimination (82%), while other aspects of performance like calibration (16%) were rarely evaluated.

Few models (13%) reported risk thresholds used for classification metrics

4 All models were at high risk of bias

All 434 models were rated as high overall risk of bias.

Main contributors were poor outcome definition, inadequate handling of missing data, potential overfitting, and limited validation.

★ Core finding ★

Current models show promise, but poor reporting, unclear outcomes, limited validation, and high risk of bias may limit clinical usefulness.

Take-home messages

- Clearly define the model prediction task and intended use
- Use reliable reference standards
- Report comprehensive evaluation, including calibration and clinical utility
- Focus on stronger methodology and reporting, not just predictive performance



Scan the QR code or visit
<https://doi.org/10.1093/jamia/cag048> to read the full paper

Affiliations

- ¹National Centre for Healthy Ageing, Monash University, Frankston, VIC, Australia
²Peninsula Clinical School, School of Translational Medicine, Monash University, Frankston, VIC, Australia
³Department of Geriatric Medicine, Peninsula Health, Mornington, VIC, Australia
⁴Aged Care Department, Austin Health, Heidelberg, VIC, Australia
⁵Department of Aged Care, Alfred Health, Caulfield, VIC, Australia
⁶School of Public Health and Preventive Medicine, Monash University, Melbourne, VIC, Australia
⁷Developmental Imaging, Murdoch Children's Research Institute, Parkville, VIC, Australia